DOCUMENT RESUME

ED 076 382 SE 015 810

AUTHOR Gullickson, Arlen R.
TITLE The Effects of Reward in Large-Scale Data Gathering.
PUB DATE Feb 73
NOTE 16p.; Paper presented at the annual meeting of the American Educational Research Association, New Orleans, Louisiana, February 1973

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Data Collection; *Questionnaires; *Research; *Statistical Surveys; *Surveys

ABSTRACT

The purpose was to test the effectiveness of a $5.00 incentive on the rate of return of mail questionnaires sent to 353 schools in the United States. The effects of the geographical region and the size of the city in which the sampled schools were located, and the school type (junior high vs. senior high) also were studied. Results showed that a $5.00 incentive did not produce actual participation from a greater percent of the sample than would be obtained if no incentive were offered, that neither the type of school nor size of city caused a differential participation rate, and that the percent of schools participating in the evaluation and providing incomplete responses varied across geographical regions. Greater actual participation once an agreement to participate had been obtained was found for schools offered the $5.00 incentive than for those offered no incentive, but the actual gain in data was not large. (DT)

FILMED FROM BEST AVAILABLE COPY

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 076382

THE EFFECTS OF REWARD IN LARGE-SCALE DATA GATHERING*

Arlen R. Gullickson

University of Minnesota

Large-scale research and evaluation studies often rely on mailed questionnaires to obtain data. However, conclusions based on such data are often subject to criticism. As Parten (1950) wrote, "Most mail questionnaires bring so few returns, and from such a highly selected population, that the findings of such surveys are almost invariably open to question." Because return rates are seldom above 40 percent, Kerlinger (1964) suggests that the researcher avoid using mail questionnaires. Nevertheless, the mailed questionnaire remains the only feasible means of contacting large numbers of persons who are scattered over a wide geographic area. Consequently, many incentive techniques have been tried in attempts to improve the response rates.

Variables for which published research data are available include: the effectiveness of including a stamped return envelope (Scott, 1961, Cullahorn, 1963), official sponsorship (Scott, 1961), length of questionnaire (Orr & Neyman, 1963), class of mail, i.e., first or third, by which the questionnaire was sent (Cullahorn & Cullahorn, 1963), color of questionnaire and inclusion of novelty items with the questionnaire (Fuqua, Nelson & Wheeler, 1970), and personally typed vs. mimeographed cover letters (Andreasen, 1970, Simon, 1967).

*This research is part of a National Science Foundation grant to the University of Minnesota to study the NSF Comprehensive Program for Teacher Education. Wayne W. Welch, Project Director.

SE 015 810

Most research on monetary incentives has been done for business and marketing concerns. Although incentives of less than 25 cents appear to be ineffective (Kephart & Bressler, 1958, Watruba, 1966, Bevis, 1948), the 25 cent incentive did produce a significantly larger response (Kephart & Bressler, 1958) than no incentive. It is interesting to note, however, in that study the 25 cents was not better than a follow-up letter. Watson (1965) in a mail questionnaire study for Business Week obtained a 48 percent return rate when he included 25 cents with a short questionnaire, and obtained a 46 percent return rate when he used a follow-up letter. By combining the two, i.e., putting 25 cents in with the questionnaire then sending a follow-up letter two days later, he achieved an 84 percent return rate. No study was found which used a questionnaire that required a half hour or more to complete and none included a large incentive, for example $5.00.

The National Science Foundation evaluation, within which this study took place, used mailed questionnaires to obtain some of its data, and incorporated a number of the above techniques in an attempt to obtain an adequate response rate. Because mailed instruments were to be sent to persons who could gain nothing by their participation, and because participants would be required to spend considerable time completing several instruments, the use of a monetary incentive appeared particularly appropriate.

It was proposed that a monetary incentive might: (1) induce a greater number of schools to agree to participate; (2) produce a higher response rate (actual participation rate) from those schools

that agreed to participate; and (3) affect the quality of returned instruments, i.e., result in higher percentage of returned packets with all instruments properly completed. However, the anticipated expense of a substantial incentive was large and evidence in support of such an incentive was meager. Consequently, the decision was made to test the effectiveness of a $5 incentive before putting the technique to extensive use.

Design and Sampling

In addition to the effect of a $5 incentive, the general NSF evaluation design provided an opportunity to study the effects of three other variables: the geographical region of the United States within which sampled schools were located; the size of the city within which sampled schools were located; and the type of school sampled. These variables were combined into a 2 x 5 x 3 x 2 fully crossed factorial design.

The monetary incentive involved an offer of $5 each to the principal and teacher of a random one-half of the schools (the experimental group) and offering no incentive to the other half (the control group). The experimental group was treated differently from the control group in two respects. First, contained in the initial letter to schools of the experimental group was the sentence, "We are offering $5 to each principal and each teacher who assists in this study." Second, two $5 checks, one to the principal and one to the teacher, were enclosed in each instrument packet sent to the experimental group schools.

The number of levels and the characteristics of the other three factors, region, city size, and type of school, were determined by the needs of the NSF evaluation. The five regions included: (1) Southern California, (2) Southern Indiana and all but the Southwestern part of Michigan, (3) all of Alabama except Birmingham, (4) all of North Dakota and most of Minnesota and Iowa, and (5) about one-half of each of the states of Nebraska, Colorado, Utah, Idaho, and Montana. (Welch & Gullickson, 1972, provides a more definitive breakdown of the sample.) Two types of schools were involved, junior high schools and senior high schools. City size was stratified into three categories: (1) cities of population greater than 50,000, (2) cities larger than 10,000 but less than 50,000, and (3) cities and towns with populations less than 10,000.

Procedure

The evaluation concerns which dictated the number of schools selected for each factor and each strata are described elsewhere (see Gullickson & Welch, 1972). To sample schools, a method called systematic sampling, (Cochran, 1953) was used. The technique of systematic sampling, though not an entirely random procedure, does produce a sample with the characteristics of a random sample when the primary sampling units are listed in a nonsystematic manner, as was done in this case. Altogether, 353 schools were sampled.

A letter requesting the participation of the principal, a teacher, and one class of students was sent to each school in the sample. The letter stipulated that the instruments to be sent would require approximately 45 minutes per person to complete, was personally addressed to the school's principal, was on NSF stationery, and was hand-signed by

the Director of the NSF Academic Year Study Program. Two weeks later, a follow-up reminder letter with a second postcard was sent to all principals who had not yet responded.

Instrument packets were then mailed to all participating schools. The mailed instrument packets contained a set of instruments, specific directions for handling the instruments, and a stamped addressed Jiffy bag for return of the completed instruments. Each teacher and administrator was asked to complete a questionnaire, the Science Process Inventory (Welch, 1966) and the Science Attitude Inventory (Moore, 1967). The teacher was also asked to administer a set of attitude and achievement instruments to a designated class.

Four weeks after the packets were mailed, a reminder letter was sent to all schools who had not as yet returned the completed packets. Enclosed with the reminder letter was a stamped postcard to be returned in the event the principal had not received the instrument packet. Each principal that had still not returned the completed instruments two and one-half weeks after the reminder letters were sent was then telephoned and asked to promptly complete and return the information.

Results

Analysis of the data by UMST 570 (Anderson & Frisch, 1971) allowed an anova solution of the four independent variables. Specifically a weighted means (least squares) analysis of city size and an unweighted means analysis of Incentive, Regions, and Schools was carried out. First, data across all schools sampled was analyzed for the two dependent variables, agreement to participate and actual participation. As can be seen in Table 3, effects for Region on the dependent variable, Actual Participation, and A x B x C x D effects for both dependent variables were obtained.

Second, the dependent variables, Actual Participation and Incomplete Response, were analyzed for the group of schools that agreed to participate. The results in Table 4 show: (1) an incentive effect as well as several interaction effects for Actual Participation, and (2) a Region effect and a Region by city size interaction effect for Incomplete Response. Hsu & Feldt (1969) and Lunney (1970) tested the applicability of anova for dichotomous data, and both indicate the anova is robust, particularly where the sample size per treatment level is large (n>50). Both studies apply to designs having equal cell sizes and Hsu and Feldt caution that factors in designs having unequal cell sizes would be subject to significance level concerns caused by possible heterogeneity of variances.

That caution seems well advised when interpreting this study's interaction effects. All significant interactions involved widely different n's and because they all included the Region variable, n's of much less than 50 occurred in every such interaction. Graphical analyses of the interactions suggest that the effects appear to be due to small differences across the five regions rather than to trends of any kind. Mean data is provided in Table 5 so that persons desiring may reach their own conclusions regarding the two way interactions.

Insert Tables 1-5 about here

This study's results do not support the hypothesis that a monetary incentive will improve mailed questionnaire response rates. There was no indication that a promise of $5 caused persons to agree to participate where they otherwise would not have done so (see Table 3). In fact, obtained mean values in Table 1 suggest the opposite. Any beneficial differences

TABLE 1

Mean Values for Agreement to Participate and Actual Participation for All Schools Sampled

Source of Variation	Sample Size n	Agreement to Participate $\bar{X}$	Actual Participation $\bar{Y}$
Region			
Southern California	66	.62	.59
Indiana-Michigan	70	.67	.67
Alabama	74	.54	.50
Midwest	74	.77	.73
Rocky Mountain Area	69	.72	.72
School			
Junior High	142	.63	.61
Senior High	211	.69	.66
City Strata			
1	91	.62	.60
2	79	.75	.72
3	183	.66	.63
Incentive			
Yes	185	.64	.64
No	168	.69	.65

TABLE 2

Mean Values for Actual Participation and Incomplete Response for Sampled Schools that Agreed to Participate

Source of Variation	Sample Size n	Actual Participation $\bar{X}$	Incomplete Response $\bar{Y}$
Region			
Southern California	41	.95	.34
Indiana-Michigan	47	1.00	.40
Alabama	40	.93	.20
Midwest	57	.95	.28
Rocky Mountain Area	50	1.00	.44
School			
Junior High	90	.97	.38
Senior High	145	.97	.31
City Strata			
1	56	.98	.39
2	59	.97	.25
3	120	.96	.35
Incentive			
Yes	119	.99	.29
No	116	.94	.39

TABLE 3

Analysis of Variance of Rate of Agreement to Participate and Rate of Actual Participation for All Schools Sampled

Source of Variation	df	Agreement To Participate MS	Agreement To Participate F	Actual Participation MS	Actual Participation F
A (Incentive)	1	.322	1.49	.121	.55
B (Region)	4	.469	2.16	.600	2.70*
C (City size)	2	.513	2.37	.390	1.75
D (Schools)	1	.148	.68	.187	.84
A x B	4	.288	1.33	.297	1.34
A x C	2	.019	.09	.020	.09
A x D	1	.223	1.03	.399	1.79
B x C	8	.053	.24	.099	.44
B x D	4	.052	.24	.026	.12
C x D	2	.185	.85	.183	.82
A x B x C	8	.321	1.48	.392	1.76
A x B x D	4	.119	.55	.222	1.00
A x C x D	2	.040	.18	.114	.51
B x C x D	8	.140	.65	.099	.44
A x B x C x D	8	.497	2.29*	.496	2.23*
Error	293			.223	

*$p<.05$

TABLE 4

Analysis of Variance of Rate of Actual Participation and Rate of Incomplete Response for Sampled Schools that Agreed to Participate

Source of Variation	df	Actual Participation		Incomplete Response	
		MS	F	MS	F
A (Incentive)	1	.132	4.43*	.428	2.08
B (Region)	4	.067	2.26	.563	2.74*
C (City Size)	2	.025	.84	.145	.71
D (Schools)	1	.016	.52	.033	.16
A x B	4	.088	2.95*	.189	.92
A x C	2	.081	2.72	.071	.35
A x D	1	.003	.09	.213	1.04
B x C	8	.026	.86	.623	3.03**
B x D	4	.075	2.51*	.035	.17
C x D	2	.008	.28	.143	.70
A x B x C	8	.017	.57	.043	.21
A x B x D	4	.123	4.13**	.461	2.24
A x C x D	2	.038	1.27	.327	1.59
B x C x D	8	.066	2.21*	.076	.37
A x B x C x D	8	.313	10.5**	.387	1.88
Error	175				

*$p<.05$

**$p<.01$

TABLE 5

Mean Values of the Significant A x R, D x B, and C x B Interactions for Sampled Schools that Agreed to Participate

Actual Participation

Region	Treatment				School			
	Incentive		No Incentive		Jr. High		Sr. High	
	n	Mean	n	Mean	n	Mean	n	Mean
Southern California	19	1.00	22	.91	18	1.00	23	.91
Indiana - Michigan	21	1.00	26	1.00	24	1.00	23	1.00
Alabama	21	1.00	19	.84	13	.85	27	.96
Midwest	29	.97	28	.93	18	.94	39	.95
Rocky Mountain Area	29	1.00	21	1.00	17	1.00	33	1.00

Incomplete Response

Region	City Size					
	One		Two		Three	
	n	Mean	n	Mean	n	Mean
Southern California	12	.33	12	.42	17	.29
Indiana - Michigan	16	.62	11	0.00	20	.45
Alabama	3	0.00	13	.15	24	.25
Midwest	9	.11	9	.56	39	.26
Rocky Mountain Area	16	.44	14	.21	20	.60

due to the incentive must have occurred at the time or after the participants received the money (see Tables 2 and 4).

Results as recorded in Table 3 lead to these conclusions: (1) A $5 incentive cannot be expected to produce actual participation from a greater percent of the sample than would be obtained if no incentive were offered, (2) neither the type of school (junior vs. senior high) nor the school setting (size of the city within which a school is located) can be expected to cause a differential participation rate, and (3) if schools are sampled from the five regions described here, the percent of schools participating in an evaluation can be expected to vary across geographical regions. Also, as can be seen in Table 4 the percent of schools providing incomplete responses can be expected to vary across regions. The study provides no answers to why the regions differ, but the differences appear great enough (Table 1) to warrant consideration if sampling is to be carried out across those regions of the United States.

Tables 2 and 4 summarize participation information of schools that agreed to participate. The Incentive Effect for Actual Participation (Table 4) establishes that greater school participation can be expected where $5 per teacher and principal participant is given, than where no incentive has been proffered. Such a conclusion suggests an incentive may better insure actual participation once an agreement to participate has been obtained. However, the actual gain in data was not large (see Table 2) and suggests that the benefits of increased data may not be worth the cost. Certainly, the NSF evaluation project will not incorporate a $5 incentive in its designs for future data gathering.

Two additional results, available but not readily apparent in Table 2, may have practical implications for large-scale data gathering. A concern of the evaluation was to obtain a large amount of information per school. As a result, the packets to be sent were bulky, expensive to put together, and expensive to mail (total postal charges averaged over $2 per packet). Because of the expense of mailing packets, the evaluation team first wrote and asked schools for their agreement to participate. It was hoped but not known that most who responded affirmatively would actually participate when they received the packet. As can be seen from Table 2, 97 percent of those agreeing did participate. That figure bodes well for others who are constrained to ask before they mail.

A second concern evidenced in Table 2 is the high incidence of partial response. In nearly a third of the participating schools, either the principal or the teacher failed to complete all assigned instruments (all student instruments were completed). The evaluation team resolved the problem by returning uncompleted instruments to the responsible persons. The importance of their full participation was explained and they were asked to complete the necessary instrument(s) and return it via the enclosed stamped envelope.

It is not difficult to visualize the seriousness of the problem. If the instruments could not (or would not) have been returned for completion, the large amount of missing data coupled with an overall 65 percent return rate would have made it impossible to reach more than very tentative conclusions regarding the regions sampled. Certainly, these results show that if much data is to be collected from each unit sampled, extra precautions

must be taken to assure that all instruments will initially be returned properly completed. Failing that, resources and schedules must be planned so that back-up measures can be taken to assure that all participants properly complete the necessary instruments.

It may be argued that other incentives, such as NSF's active support of the evaluation, masked the positive effect of the monetary incentive. That explanation would be more attractive if the participation rate had been quite high, say 85 percent to 90 percent. However, even if it were true, the argument provides far better support for incentives other than money than it does for the monetary incentive.

It is the opinion of this writer, that monetary incentives show little promise for being a cost effective method of data collection via mailed questionnaires. Certainly the problem of how to obtain a high response rate from mailed questionnaires has not been resolved. However, unless and until a viable alternative is obtained, mailed questionnaires will remain a necessity for many studies. Therefore, potential solutions must be investigated. Possibly an investigation of the effectiveness of preliminary contact(s) with subjects coupled with stringent follow-up procedures would be an appropriate direction to turn.

BIBLIOGRAPHY

Anderson, D. & Frisch, M., "UMST 570 Multivariate Analyses of Variance," UMST Computer Programs Manual, Minneapolis: University Computer Center, University of Minnesota, 1971.

Andreasen, A. R., "Personalizing Mail Questionnaire Correspondence," Public Opinion Quarterly, 34, 1970, pp. 273-277.

Bevis, J. C., "Economic Incentives Used for Mail Questionnaires," Public Opinion Quarterly, 12, Fall, 1948, pp. 492-493.

Cochran, W. G., Sampling Techniques, New York: John Wiley, 1963.

Gullahorn, J. E. & Gullahorn, J. T., "An Investigation of the Effects of Three Factors on Response to Mail Questionnaires," Public Opinion Quarterly, 27, 1963, pp. 294-296.

Gullickson, A. R. & Welch, W. W., "Applying Experimental Designs to Large-Scale Program Evaluation," University of Minnesota, Minneapolis (Mimeo), 1972.

Hsu, T. & Feldt, L. S., "The Effect of Limitations on the Number of Criterion Score Values on the Significance Level of the F-test," American Educational Research Journal, November, 1969, pp. 515-527.

Kephart, W. M. & Bressler, M., "Increasing the Response to Mail Questionnaires: A Research Study," Public Opinion Quarterly, 22, Summer, 1958, pp. 123-132.

Kerlinger, F., Foundations of Behavioral Research, New York: Holt, Rinehart & Winston, Inc., 1964.

Lunney, G. H., "Using Analysis of Variance with a Dichotomous Dependent Variable: An Empirical Study," Journal of Educational Measurement, 7, 1970, pp. 263-269.

Moore, R. W., The Development, Field Test and Validation of the Scientific Attitude Inventory, Unpublished Doctor's Dissertation, Philadelphia, Pennsylvania: Temple University, 1969.

Orr, D. B. & Neyman, C. A., "Considerations, Costs and Returns in a Large-Scale Follow-Up Study," Journal of Educational Research, 58, April, 1965, pp. 373-378.

Parten, M., Surveys, Polls, and Samples, New York: Harper & Row, Chapter 1, 1950.

Pucel, D. J., Nelson, H. F. & Wheeler, D. N., "Questionnaire Follow-Up Returns as a Function of Incentives and Responder Characteristics," Vocational Guidance Quarterly, 19 (3), March, 1971, pp. 188-193.

Scott, Christopher, "Research on Mail Surveys," Journal of the Royal Statistical Society, Series A, 124, 1961, Part 2, pp. 143-195.

Simon, R., "Response to Personal and Form Letters in Mail Survey," Journal of Advertising Research, 7 (1), 1967, pp. 28-30.

Watruba, T. R., "Monetary Inducement and Mail Questionnaire Research," Journal of Marketing Research, 3, 1966, pp. 398-400.

Watson, J. J., "Improving the Response Rate in Mail Research," Journal of Advertising Research, 5, June, 1965, pp. 48-50.

Welch, W. W., Welch Science Process Inventory (SPI), Form D, 6115 Woody Lane, Minneapolis, Minnesota 55432, 1966.

Welch, W. W., & Gullickson, A. R., "A Strategy for Evaluating the NSF Comprehensive Program for Teacher Education," School Science and Mathematics, Spring, 1973, (in press).